5Б

УДК 519.8

І.В. Баклан, Г.А. Степанкова
Національна академія управління, м. Київ, Україна
iaa@ukr.net, a.a.stepankova@gmail.com

Про деякі нові особливості використання прихованих марковських моделей для аналізу та прогнозування часових рядів

У статті розглядаються деякі нові підходи до аналізу та прогнозування часових рядів за допомогою математичного апарату різноманітних прихованих марковських моделей. Пропонується використання сумішей неоднорідних прихованих марковських моделей для аналізу часових рядів, а також для прогнозування часових рядів пропонується використання гібридної моделі – нечіткої прихованої марковської моделі.

Вступ

Метою даної роботи є розробка нових підходів до аналізу та прогнозування часових рядів на основі математичного апарату прихованих марковських моделей.

Пропонується підхід до побудови гібридної моделі на основі апарату нечітких множин та прихованих марковських моделей (ПММ), яка отримала назву нечіткої прихованої марковської моделі (НПММ). Будемо застосовувати вже відомі принципи гібридизації різноманітних моделей з прихованими марковськими моделями [1], [2]. Особливістю цього підходу є застосування алгоритму нечіткого очікування в ПММ. Цей підхід полягає у застосуванні алгоритму на кшталт алгоритму максимального очікування, подібного до алгоритму Баума-Велша.

Надалі нами розглянуті випадки дискретних та безперервних ПММ та показується вища продуктивність НПММ порівняно зі звичайними ПММ.

Кластеризація нечітких середніх (КНС) є одним з найбільш використовуваних підходів кластерного аналізу [3]. На основі цього підходу відоме рішення, що отримало назву нечіткого квантування векторів [4], але в ньому передбачений виключно дискретний варіант – спостереження дискретних символів. Він не застосовується до безперервних ПММ, тобто коли спостереження є безперервним та стоїть задача моделювання функції щільності ймовірностей.

В останні роки для аналізу часових рядів застосовується апарат прихованих марковських моделей (ПММ) [5], [6]. У цьому випадку неоднорідність може виникнути в часі таким чином, що буде зачіпати прихований процес, тобто ймовірності переходу між прихованими станами.

Формалізація моделі

Перед нами стоїть проблема створення підходу, який можна було б застосовувати як для дискретних, так і для безперервних ПММ. Відзначимо, що фактично нам треба дати відповіді на такі три нагальні питання.

Наш підхід буде своєрідною гібридною моделлю використання теорії нечітких множин та алгоритму максимізації для оцінки параметрів λ моделі ПММ. Для вирішення наших проблем звернемося до визначення функції приналежності $s \in S$, $U = [u_s(o)]$, $\sum_s u_s(o) = 1$, $0 < u_s(o) \leq 1$ для максимального $L(O,\lambda) = \log P(O|\lambda)$ на досліджуваному наборі даних O за допомогою функції, визначеної як $Q_F(U,\lambda)$, та певними зусиллями для максимізації λ. При цьому F – константа більша за 1. Також за нашим алгоритмом НМО повинно бути отримано $Q_F(\overline{U},\overline{\lambda}) \geq Q_F(U,\lambda)$, що фактично отримується у два кроки.

Перший крок – визначення $\overline{U}$, де $Q_F(\overline{U},\lambda) \geq Q_F(U,\lambda)$. А вже на другому кроці знаходимо λ, для якої справедливо співвідношення $Q_F(\overline{U},\overline{\lambda}) \geq Q_F(\overline{U},\lambda)$.

Шкала незалежних досліджень O та s-станів визначається за наступними рівняннями:

$$Q_F(U,\lambda) = -\sum_s u_s^F(o) d_s^2(o), \qquad d_s^2(o) = -\log P(o,s|\lambda).$$

Для максимізації $Q_F(U,\lambda)$ в останньому рівнянні треба максимізувати КНС функцію:

$$u_s(o) = \left| \sum \left(\frac{d_s(o)}{d_i(o)} \right)^{\frac{2}{F-1}} \right|$$

У цьому рівнянні $d_s(o)$ – відстань або шкала незалежних досліджень O та станів I, вираз отримання якого буде наведений пізніше. Оцінка параметрів λ має наступні кроки.

1. Спочатку визначимо величину константи F, визначити s та першу частину параметрів для λ.

2. Надалі обчислимо $\overline{U}$ та $Q_F(\overline{U},\lambda)$ на основі λ та $\overline{U}$.

3. Визначимо $\overline{\lambda}$, коли максимальне $Q_F(\overline{U},\lambda)$.

4. Міняємо U на $\overline{U}$ та λ на $\overline{\lambda}$ та повторюємо з другого кроку.

Тепер розглянемо дискретні нечіткі приховані марковські моделі.

Нехай у нас буде $Q_F(U,\lambda)$ – функція з наведеного вище рівняння:

$$Q_F(U,\lambda) = \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{M} u_{ijt}^F(o_t) \log P(O, s_t = i, s_{t+1} = j|\lambda),$$

де $u_{ijt}^F(o_t)$ – функція приналежності, що вказує, за якою мірою послідовності спостережень O належить стан i на момент часу t та стан j на момент часу $t+1$, та яка задовольняє наступним умовам:

$$\sum_{i=1}^{N} \sum_{j=1}^{M} u_{ijt}(o_t) = 1, \; 0 < u_{ijt}(o_t) \leq 1.$$

Нечіткі Q-функції досліджуваної послідовності даних для нечітких ПММ визначаються наступним чином:

$$u_{ijt}(o_t) = \left| \sum_{k=1}^{N} \sum_{l=1}^{M} \left(\frac{d_{ijt}(o_t)}{d_{klt}(o_t)} \right)^{\frac{2}{F-1}} \right|^{-1},$$

де $d_{ijt}(o_t)$ – відстань недосліджуваного o_t в стані i на момент часу t та стані j на момент часу $t+1$, яка задовольняє наступному співвідношенню:

$$Q_F(U,\lambda)=\sum_S u_s^F(O)\log P(O,S|\lambda).$$

Застосовуючи ті ж самі методи оптимізації, що й для традиційних ПММ, параметри нечітких ПММ оцінюються наступним чином:

$$d_{ijt}^2(o_t)=-\log P(o_t,s_t=i,s_{t+1}=j|\lambda)=-\log\{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)\},$$

$$\overline{\pi}_i=\sum_{j=1}^{N}\overline{u_{ijt}^F}(o_t),\ \overline{a}_{ij}=\frac{\sum_{i=1}^{T-1}\overline{u_{ijt}^F}(o_t)}{\sum_{i=1}^{T-1}\sum_{j=1}^{N}\overline{u_{ijt}^F}(o_t)},\ \overline{b}_i(k)=\sum_{t\in o_t=v_k}\frac{\sum_{i=1}^{T-1}\overline{u_{ijt}^F}(o_t)}{\sum_{i=1}^{T-1}\sum_{j=1}^{N}\overline{u_{ijt}^F}(o_t)}.$$

Надалі розглянемо принцип гібридизації безперервних нечітких ПММ.

Тепер нехай $Q_F(U,\lambda)$ – нечітка функція параметрів ПММ – визначається на основі НМО алгоритму:

$$Q_F(U,\lambda)=\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{k=1}^{M}u_{ikt}^F(o_t)\log[c_{ik},N(o_t,\mu_{ik},\sigma_{ik})].\tag{1}$$

Аналогічно до дискретного випадку, нехай $u_{ikt}(o_t)$ – функція приналежності, що вказує, за якою мірою послідовності спостережень O належить стан i на момент часу t та стан j на момент часу $t+1$, та яка задовольняє наступним умовам:

$$\sum_{i=1}^{N}\sum_{k=1}^{M}u_{ikt}(o_t)=1,\ 0<u_{ikt}(o_t)\le 1.$$

Елементи нечітких функцій в рівнянні (1) обчислюються наступним чином:

$$\overline{u_{ikt}}(o_t)=\left|\sum_{j=1}^{N}\sum_{l=1}^{M}\left(\frac{d_{ikt}(o_t)}{d_{jlt}(o_t)}\right)^{\frac{2}{F-1}}\right|^{-1},$$

де відстань $d_{ijt}(o)$ знаходимо зі співвідношення, наведеного нижче, та величина якого показує масштаб неспостережності o_t в стані l та момент часу t та комбінації k -х.

$$d_{ikt}^2=-\log P(o_t,s_t=j,k_t=k|\lambda)=-\log\{\alpha_t(i)a_{ij}b_j(o_{t+1})N(o_t,\mu_{jk},\sigma_{jk})\beta_{t+1}(j)\}.$$

Використовуючи той самий метод оптимізації стосовно НМО параметрів нечіткої ПММ, отримуємо наступні співвідношення оцінки:

$$\overline{c_{ik}}=\frac{\sum_{t=1}^{T}\overline{u_{ikt}^F}(o_t)}{\sum_{t=1}^{T}\sum_{j=1}^{M}\overline{u_{ijt}^F}(o_t)},\ \overline{\mu_{ik}}=\frac{\sum_{t=1}^{T}\overline{u_{ikt}^F}(o_t)o_t}{\sum_{t=1}^{T}\overline{u_{ikt}^F}(o_t)},\ \overline{\sigma_{ik}}=\frac{\sum_{t=1}^{T}\overline{u_{ikt}^F}(o_t)(o_t-\overline{\mu_{ik}})(o_t-\overline{\mu_{ik}})^T}{\sum_{t=1}^{T}\overline{u_{ikt}^F}(o_t)}$$

Ми представляємо кінцеву суміш НОПММ, де кожний з компонентів відрізняється лише прихованим процесом. Наша мета має два аспекти: класифікувати предмети за їх поведінкою у часі та за оцінками непараметричної максимальної правдоподібності змішаних НОПММ. Ефективність даної пропозиції планується перевірити за допомогою експериментів на побудованій моделі дослідження.

Багато уваги під час аналізу даних часових рядів приділяється еволюції окремих прихованих характеристик, які можуть бути оцінені непрямим способом через декілька

відповідних змінних. Моделі такого роду аналізу відомі під назвою прихованих марковських моделей (ПММ), де інтерес до еволюції прихованої характеристики представляється неспостережуваним ланцюгом Маркова з невеликою кількістю станів та відповідних змінних, які вважаються умовно незалежними від заданого прихованого процесу.

Така схема аналізу з використанням однорідних ПММ може бути застосована для аналізу даних кількох людей, які спостерігаються періодично, для опису розвитку хвороби або змін у динаміці поведінки. Але припущення про однорідність в моделі Маркова, тобто що ймовірності переходів є постійними в часі та між окремими особами, можуть бути серйозним обмеженням для використання цих моделей.

Щоб розв'язати цю проблему, пропонується ввести поняття неоднорідних ПММ. Це дозволяє матриці переходу прихованого ланцюжка змінюватися у часі або за іншим параметром на основі екзогенних коваріацій. Треба зауважити, що не завжди наявні коваріації в змозі охопити всі окремі однорідні джерела дослідження. У цьому випадку приходимо до так званого класу змішаних ПММ. Недосліджувані окремі неоднорідності можна моделювати за допомогою підходу прихованого класу. Особи групуються у класи за прихованою неконтрольованістю з урахуванням їх динамічної поведінки. Отримана модель є сумішшю ПММ, де непомітні джерела неоднорідності. Цей підхід дозволяє аналізувати дані в різних ситуаціях. Тим не менше, такий вибір може на практиці мати певні складності в обчисленні, саме ті складності, що виникають при оцінці чергового кроку. Оцінка ПММ може здійснюватися через вже добре відомі методи, такі як метод максимальної правдоподібності, метод Монте-Карло або через метод оцінки підпростору. Всі методи створюють певні проблеми із складністю обчислень. В нашому випадку змішаних ПММ може стати цікавим і цілком корисним метод непараметричної максимальної правдоподібності (НПМП), де розподіл випадкових впливів оцінюється дискретними багатовимірними випадковими величинами з кінцевим числом опорних точок. У результаті НПМП оцінки призводять до кінцевої суміші ПММ, як і у випадку прихованого класу ПММ.

Висновки

Підвищення ефективності оцінки ПММ можна продемонструвати на основі розпізнавання траєкторій, які залишає «мишка». Навчання відбувалося за 8 траєкторіями (сеансами роботи за чітко визначеним алгоритмом), які виконували 8 людей. Для перекодування використовували структурний лінгвістичний підхід [7], [8].

При цьому якість розпізнавання на проведених дослідах показала 10% помилок при використанні традиційного підходу для дискретних ПММ, на відміну до майже 5% при застосуванні нечітких дискретних ПММ.

До того ж на безперервних ПММ були отримані 8 % для традиційного та 4 % – для використання НМО.

Наведені дані підтверджують підвищення якості розпізнавання для цього типу сигналів.

Ми запропонували використання кінцевої суміші НОПММ за рахунок розширення теорії Діаса для неоднорідного випадку [9]. Передбачається, що недосліджувана однорідність впливає тільки на прихований процес, тоді як окремі розподіли, стала часу, випадкові параметри мають дискретний характер. В такий спосіб ми визначаємо доволі гнучку модель для опису широкого спектра динаміки поведінки, що забезпечує як неконтрольовану класифікацію суб'єктів та напівпараметричну оцінку змішаної

ПММ. Параметри моделі оцінюються за допомогою максимальної правдоподібності та за допомогою ЕМ-алгоритму.

Такий підхід можна розглядати як можливе вирішення питань, які виникають при проведенні обчислювальних експериментів зі змішаними НОПММ.

Література

1. Баклан І.В. Гібридні технології в проектуванні інтелектуальних систем прийняття рішень / І.В. Баклан // Сучасні інформаційні та інноваційні технології на транспорті : матеріали Міжнародної науково-практичної конференції. – Херсон : Видавництво Херсонського державного морського інституту, 2009. – Том 1. – С.32-37.
2. Баклан І.В. Основні проблеми при застосуванні прихованих марковських моделей / І.В. Баклан, Г.А. Степанкова // Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту : матеріали міжнародної наукової конференції. – Херсон : ХНТУ, 2009. – Том 2. – С. 430-432.
3. Jamse C. Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms / Jamse C. Bezdek. – Plenum Press, New York and London, 1987.
4. Tsuboro E. On the Fuzzy Vector Quantization based Hidden Markov Models / E. Tsuboro, J. Nakanashi // Proc. Of IEEE Intl Conf. ASSP. – 1994. – C. 637-640.
5. Баклан І.В. Імовірнісні моделі для аналізу та прогнозування часових рядів / І.В. Баклан, Г.А. Степанкова // Искусственный интеллект. – 2008. – № 3. – С. 505-515.
6. Степанкова Г.А. Деякі підходи до навчання прихованих марківських моделей / Г.А. Степанкова, І.В. Баклан // Матеріали міжнародної наукової конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту». – Євпаторія. – 2008. – Т. 3 (ч. 2). – С. 87-91.
7. Баклан І.В. Структурний підхід до аналізу та прогнозування поведінки часових рядів / І.В. Баклан, Ю.М. Селін // Вісник Херсонського національного технічного університету. – Херсон : ХНТУ, 2005. – № 2. – С. 27-31.
8. Баклан И.В. Анализ поведения экономических часовых рядов с использованием структурных подходов / И.В. Баклан, Ю.Н. Селин // Вісник Херсонського національного технічного університету. – Херсон : ХНТУ, 2006. – № 2. – С. 29-34.
9. Dias J.G., Vermunt J.K., Ramos S. Mixture hidden Markov models in finance research // In: Fink A., Berthold L., Seidel W. and Ultsch A. (eds.) Advances in data analysis, data handling and business intelligence, Springer: Berlin-Heidelberg, 2010. – P. 451-459.

И.В. Баклан, А.А. Степанкова

Про некоторые новые особенности использования скрытых марковских моделей для анализа и прогнозирования временных рядов

В статье рассматриваются некоторые новые подходы к анализу и прогнозированию временных рядов с помощью математического аппарата скрытых марковских моделей. Предлагается использование смесей неоднородных скрытых марковских моделей для анализа временных рядов, а также для прогнозирования временных рядов предлагается использование гибридной модели – нечеткой скрытой марковской модели.

I.V. Baklan, G.F. Stepankova

On Some New Pecularities of Hidden Markov's Model Usage for Analysis and Prognosis of Time Series

In this paper we consider some new approaches to analysis and prognosis of time series by means of hidden Markov's models. We propose for time series analysis to use mixtures of unhomogenous hidden Markov's models. Also we consider some properties of hybrid fuzzy hidden Markov's models for time series prognosis.

Стаття надійшла до редакції 30.06.2010.